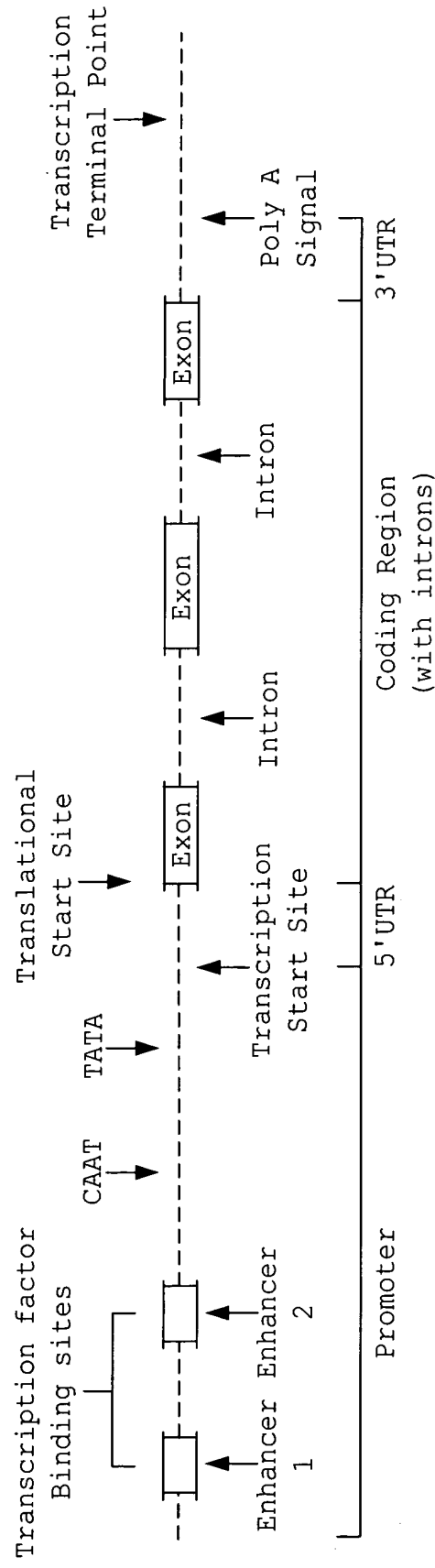


FIG. 1

SCHEMATIC OF A GENE



Sequences/motifs that specific DNA conformation, chromatin conformation, extent and position of base methylation and binding sites of proteins that control of these.

Gene

FIG. 2

-4242	GCATGCACTG	CCACAAGTAG	TGAACATCATG	GTTTTACCTC	CTCAAGTAGA
-4192	AAACCTTTTG	AGTGAATTTG	AAGATTTATT	CTCCCAAGAA	GGACCCATTG
-4142	GGCTTCCTCC	TCTTAGGGGG	ATAGAACATC	AAATTGACTT	TATACCGGGG
-4092	GCAAGCCTAC	CAAATAGGCC	TCCTTATAGA	ACCAACCCCG	AGGAAACAAA
-4042	GGAGATAGAA	TCACAAGTTC	AAGACTTGTT	GGAGAAGGGT	TGGGTTCAAA
-3992	AGAGCCTAAG	CCCTTGTGCT	GTACCTGTCT	TGTTGGTGCC	AAAAAAAGAT
-3942	GGAAAATGGC	GTATGTGTTG	TGATTGTAGA	GCAATCAACA	ACATCACCAT
-3892	CAAGTATAGG	CATCCAATCC	CAAGGCTTGA	CGATATGCTT	GATGAATTGC
-3842	ATGGGTCAAC	TCTATTCTCC	AAAATTGACC	TTAAAAGTGG	ATATCACCAA
-3792	ATTCTGAATCA	AGGAGGGTGA	TGAGTGGA	ACCGCTTTTA	AGACCAAATT
-3742	TGGATTATAT	GAGTGTTGG	TGATGCCCTT	TGGTCTTACT	AACGCTCCAA
-3692	GTACATTCAT	GAGGCTTATG	AATCACACCT	TGAGGGATTG	TATAGGTAAA
-3642	TATGTAGTAG	TTTATTTTGA	TGATATCTTA	GTATATAGTA	AAACCCTAGA
-3592	AGACCATCTA	AGTCACCTTA	GGGAAGTTCT	TCTAGTTCTT	AGGAAAAATA
-3542	GTCTTTTTGC	CAATAGGGAT	AAGTGTACCT	TTTGTGTAGA	TAGCGTAGTC
-3492	TTTTTAGGCT	TTATAGTAAA	CCAAAAGGGG	GTGCATGTAG	ATCCCGAGAA
-3442	AATCAAAGCC	ATCCGCGAGT	GGCCAACCTC	ACAAAATGTA	AGTGATGTGA
-3392	GAAGTTTTCA	TGGGTTAGCT	AGCTTCTATA	GAAGGTTTGT	TCCCAATTTT
-3342	TCTAGCCTAG	CTTCTCCCTT	GAATGAACCT	GTAAAAAAG	ATGTTGCATT
-3292	TTGTTGGAAT	GAAAAGCATG	AGCAAGCCTT	TCAAAGGCTA	AAAGCTCACT
-3242	CACCAATGCA	CCCATCCTAT	CTCTTCCAAA	TTTTTCCAAA	CTTTTGGAGA
-3192	TAGAGTGTGA	TGCATCGGGA	GTAGGCATAG	TGCGGTTTTG	TTGCAAGGTG
-3142	GACACCCCTT	GCTTATTTTA	GTGAAAACT	CCATGGTGCC	ACCCTCACTA
-3092	CCCCACCTAT	GACAAAGACT	CTATGCTCTT	GTGCGACCCT	AAAGACTTGG
-3042	GGAACACTAC	CTTGnGTCCC	AAAGAATTTG	GnTATCCATA	GTGATCACGA
-2992	GTCTTTTAAA	TATTTAAAGG	GCCAACACAA	GCTCAATAAG	AGACATGCTA
-2942	AATGGATGGA	ATTTCTTGAA	CAATTTCTT	ATGTCATCAA	ATACAAGAAA
-2892	GGGAGCACCA	ATATAGTGGC	CGATGCTCTT	TCTAGACGGC	ACACTCTCTT
-2842	TTCAAAACTA	GGTGCCCAAA	TTCTTGGA	TGACCACATA	AGAGAGCTTT
-2792	ATCAAGAAGA	TCAAGAACTC	TCATCCATCT	ATGCCCAATG	TCTACATAGA
-2742	GCACAAGGAG	GTTACTATGT	GTCCGAGGGA	TATCTTTTTA	AAGAAGGAAA
-2692	ACTTTGCATT	CCCCAAGGAA	CACATAGAAA	ACTCCTTGTC	AAAGAATCAC
-2642	ATGAAGGGGG	ACTCATGGGC	CATTTTGGAG	TTGATAAAAC	TCTAGACTTT
-2592	TAAAAGCAAA	ATTTTGTTGG	CCACACATGA	GGAAAGATGT	CCACGACATT
-2542	GTCTAGAGTA	TCTCATGTTT	AAAAGCAAAG	TCTAGAACAA	TGCCGCTGGA
-2492	CTCTACACCC	CTTTGCCGAT	TGCAAAGCTC	CTTGTGAAGA	CATTAGCATG
-2442	GATTTTCATTT	TAGGACTTCC	TAGGACTGCA	AGAGGCCATG	ACTCTATCTT
-2392	TGTGGTAGTG	GACCGTTTTA	GCAAAATGTC	TCACTTTATT	CCATGCCACA
-2342	AAGTAGATGA	TGCTCAAAAT	ATTTCTAAAC	TCTTCTTTAG	AGAAGTGGTG
-2292	AGACTCCATG	GTCTCCCTAG	AAGTATAGTG	TCCGATAGAG	ATCACCTTAA
-2242	ATATATAATT	ATACACTTGT	TTTTTTTCTC	TTTTTTTATTT	TATCAAGTAA
-2192	AAAGTATTTG	TTCTAGATTA	TTATGAGTAT	ATACTTACTT	TCTGTATTTT
-2142	ATTTCTTTCT	ATTTTTTATG	ACGATGAAAT	TTCTTATTAT	ATCCAGACTT
-2092	TTCATATATA	TTTTTTATTT	TTTTCCATCT	AGATGCTCTG	TACTTTTCTT
-2042	CAGTTGAAAT	TTCCACTCTC	CAACAAAACA	TCATTCAAGT	TTTGTATAAC
-1992	ACTGTGACGT	TAACCAGTTA	AAATAAGAAA	ATCATGTAAT	ATAAATTATT
-1942	TCAGTAGATA	TTTTTAGAATT	ACAAATACGA	TAAATAATTA	AATTTAAAAA
-1892	ATTATTAAAC	AATGAATTTT	TTTGAAATTT	AATATAAAAC	TTAGACTTGT

FIG. 2 (Continued)

-1792 ACATCGAATT TGGGTGCTTT ATGCCGCTTT ATCTTCATCT GCACCTTCAA
-1742 ATTAATAATT TAATTCCGGA AAATAATAAA CCCACACACT GTTTTATGCA
-1692 TATATTAAGA TAAATAAAAG AGAACTATTT TAAAGAATAT AAAATAATAA
-1642 ATGTAACAAA TGATGTCAC TAAAGAAGAAA AAAATTAACA AGAATTGTAA
-1592 TATATTTCTT TATGAAATGT TTTGTGCATT ACCGAGAGAG GTCGAACATG
-1542 ATACACGCAA GCATCTAACT AGTTTGGTAA TTCCTTTTCA ACATCGnTAA
-1492 GCACATCACA CTAAATTTAC TTTAAATAGA TAAATTAGAT TCAATTGGAT
-1442 GACATTAATT TATAATACTC TATCCAAAAT TATAACTATA AATAAAAAGT
-1392 TATTTTTAGA AAATAAGTAA TGAAAATTTA ATTCTAAAAT TTATAACACT
-1342 TTTATGCTGT GTTTGTTTCG AAGCATAGAA AAATAAAAAG TTATTGTTGG
-1292 GAATGAAAAG TGAAGAAAAT CATGTAATAA AAACAAAATG ACACGACAAT
-1242 CAAAAAAGAA GTTTTCATGC AAAACTTTTT TCAAATTTA CACTTTTATG
-1192 ATGTGTTTGT TTCGAAGTGT AGAAAAACGA AAAGTTATTA TTGGTAATGA
-1142 AAAGCGAAGA AAATCACGTA ATAAAAACAA AGCAAGATGG CACGACAATC
-1092 AAAAAAAGT TTCTACACAA AACTTTATTC AAAATTTACA ACACTTTTAT
-1042 GTTGTGTTT GTTTCCGAGG TATAGAAAAA CAAAGAATTA GTGTTGGTAA
-992 TGAAAAGTGA AGAAAACCAT GTAATGAAAA CAAAATGGCA CGACAATCAA
-942 AAAAAGTTTT CACGCAAAAT TTTCTTCAAA ATTTATAACA TTTTCATGTT
-892 GTGTTTGTGTT CAAAGCCTAG AAAAACGAAG AGTTACTATT GGTAATGAAA
-842 AGCGAAGAAA ACCACATAAT AAAAAACAAA TGGCACGACA ATCAAGAAAA
-792 AGTTTTTACA CAAAACTTTT TTCAAATTT ACTATGTTTA TTTCGAAATT
-742 TAGAAAAACG AAGAGTTATT ATTAGTAATG AAAAGCGAAG AAAACTACGT
-692 AATAAAAAAC AAAATGGCAC GACAATAAAA AAAGTTTTCA CGCAAATTT
-642 TCTTGGTGCG CAGAAAGTTA TATATATTAA TTAATTAATT TTCATTTACT
-592 TTTTTCCCTT TTTATTTTAA AGTTAAATTA TTATTATTTT CATTTAAAT
-542 ATAAATATTA TTTAAATATA AAAAATATAA CCTTAATCAA AACAAAGCCT
-492 TAATCTAAAA TTTACAACAC TTTTAACCTT AAAATTAAT TTAAGGAA
-442 AATGATAGTG TGACAACTAA AAAAGTTGTA TACAACCCTG TCATAGGTTT
-392 AGAAATAAAT ATATATAATA AAGAGTAAAT TTGTAATTAA ATGATATAAA
-342 AAAGTATTAA AATAATAATA TTTAGAGTAG TAATATGGTT GTATAAAAAA
-292 ATGTGGTTGT CCATATATCA TTATTTACTT TAAAATATCA TGACAAATAT
-242 TTTACCCGAA AGATGGAAAG AACGAAAAGA GCGTTGGATA ATGGAAAAAT
-192 ACAAGCAATC TCCCTCCAGT ACTTTGCATA ACATTTTGTA TTAGTGATGA
-142 GTTTTTTATC ATATATATTT AGAATATAGG AAAATTTTAG AATCACGTGG
-92 ATAGCTATAT AATAGTAATA TTTTAATTTA TAATGTAGTT GATTTTATTT
-42 GTCAACTGGT ATACATAAAT ATGTGTTGAT AGTGGGTGAC TTGTGGCTTA
9 AAGAAATGTC CAGAGGCTGA CAACAACCTCT GCACAGACTA GCGTAAAC
57 **ATG** AAG TCC AAT TTT GCT ATT TTC GTA GTC TTT TCT CTT CTT CTT
1 M K S N F A I F V V F S L L L
102 CTG GTACCTCTTCAATCTTCTCTACAAAACTCTGTTGCTCTTTCACCTCTGTTTGTA
16 L
160 ATTTTGTTTACACTTTTGGAAAATTGAAGCTGATATATATGTAACAACCTTTTCAGTTTT
219 GTCTGCACTGAACTGATAGAAAAATATACGTTTTGTGGATATATATAG GTT GGC
17 V G
274 AGT TGC AGC TGC GCA AGA AAA GAC ATG AGA GGG TAT TGG AAG GAT
19 S C S C A R K D M R G Y W K D
319 ATG ATG AAG GAG CAA CCT ATG CCA GAA GCA ATC AAA GAC CTT ATT
34 M M K E Q P M P E A I K D L I

FIG. 2 (Continued)

	GAG	GAT	TCA	GAA	GAA	GTG	TCA	GAA	GCA	GGG	AAG	GGT	CGT	TTT	GTT
49	E	D	S	E	E	V	S	E	A	G	K	G	R	F	V
	AGG	GAC	TTC	GAT	GTA	AAG	CCT	AAT	GTC	ATA	TTA	TAT	CAC	ACA	CAT
64	R	D	F	D	V	K	P	N	V	I	L	Y	H	T	H
	GTT	GTG	CCC	ATG	AAG	CAG	AGG	CAG	AAG	AAT	AAA	GAT	TGA		
79	V	V	P	M	K	Q	R	Q	K	N	K	D	.		
493	AGACTATGTGATTGGCAGTTTCAGACTTATTTGGCACCAAATTTATGATGCTCTTGTTGC														
553	TGTTTTCAA AATTTGTACTCAA ACTTTGAACCCTTTGCAGCATCTTGCTTCTTTTTGGTCT														
613	TGCTGAATTTTGT CACAGTTATACTGTCACGAATAGTTTCTCTTCATAATAAGCAACTTT														
673	TCCTCTC														

FIG. 3

101001 CAAAACAAAAGCAAATGCCGGTTTTCTTATTATTATTTTCGAACTTTAGAC
100151 CTTTTTGTAACGTTTCTTTAATTTTTTTCCTTGATAAAGAACCCTATTAT
100201 ATCTTAGCTAAATATTTACCTCATTTTGTATTATGAGCTAAACCACCCCAA
100251 AAATATTGTAGTTTTGCTTTCGGATTTAACTGCCAAGCAAGTGATTAGAT
100301 ATATTAAAGGAAAATGAATGAAAGGACAAAAAATATAAACGACAATATT
100351 TGAATACTGATATTTATCTCCATTCTCAAATATTTTTGATTTATTGTGAC
100401 AATATTTGGTTGTTTCCCATTGCTACATCTTTGAGGACATGAAATGATA
100451 ACATATATATGAACGAGTATAATACATTCTCGTTTCATTTTACAAATAAT
100501 GTCATTTTATGCTAACATTTTTTATTTAAAAATTATCCTTATAAGATTTT
100551 AGTGTATTATTTTACCATGGTACTGTAAAGTCGGATGCTATATATATATA
100601 TATATATATATATATCAAAAATGACACTGAAGAATTTATTTGAACTAAAA
100651 CTAAAAACGTAAAAATAAAAAGAATTTTTCAAAAATCAAAAATTTTATATA
100701 AAAATATAGATAAAAATGTTAATATAGTACAACTTCTATTCAAACAGAGAG
100751 AATAAATCTTCTATAGACAGTGAATATCCATTATAATAACGAGCAATAGT
100801 TGTAATGTTGCAGTACAAAAAGAGAATTGTAATATTTGTGCATGATTGAG
100851 AAATCTAAGTTGACTTTGAATTAAAAGGCTAATTCCAACAAGTACATGTA
100901 GAAGTTGACTATAGCTATATATTTACTACAAATTGATCATTTCAAGAAAG
100951 ACATTTAAATTAAGATATGCATGCATGACTTGATTGAACCCCACTCGCTT
101001 GCTTCGTGCCATTTCGACAAGATGTTACTTTTAAATGCAAGGTAAATTATG
101051 GATATACTCTTCTGTATTTTTTTGTAGTAGATATTTTTACGAAAATTGTTT
101101 TTTTTCCAAAATCAAATGATATTTATTAATTTTCAATATAGAATTAATTA
101151 AATTTTAATTAATTTTGAAGATTTATATGCTGCAGATTAGATTACCATTG
101201 GTGAAATCATGTTTAGGTAAATAATAAATGATGTTGTAGTTTAGGAAAAA
101251 AAAAAATTCTTTAATCTTTATGTAAGAATGTTAACTTCAATTATAAAAA
101301 TATGAAGCAGTATTATATAAGATGTTTAACTAATCGAATAATATTTTTTG
101351 GGATGAAATTTTCTTGCATATGTTTCTAAAAAATAATATGTGAAAAATT
101401 AACATTCATTGTATGTTTATAAGAAATATATGTGAGTTTGTTTAGATAA
101451 ATAATACTTAAAATTAAGAATTTGTAAAGTTATACTGCACTTCAAATATG
101501 TTATTTTTTTCCTTTTATTTTAAAATATCAGCAACATTCTAAATGATTTTAT
101551 TTTCTTTAAAAAATTGAAAAAATGAAATTAGCAAATATGTAAAATTTAAA
101601 ACGAATTTAAGAAAAAACTTTGTAAAGATATGATATGCTTTATAAAAAAA
101651 ACTTGGTGGCGTACCTACTAAATATGATCACATTAGAGATTTGTATCCTT
101701 TAGCATATAGTATGTAGTATAGATATCTATATTTTTTATTTATTAAAGAGC
101751 ATATTCATAATATAGGTATTATATGTTAATTACAATAAACGTTCAATTCTG
101801 TTATGTTAGTTTTTGAAGAACTTATTGCGTGTGCATATCAATGTGAGAAA
101851 GCGACTCCACATGTGAGATGTTGGTCTGAGAAAGCTTTCTGCACTTGGTC
101901 GGAACACTTCATGGACTAGAATGCAATCCATCTATTCAAAGAAAAGCAG
101951 TTGTCCATGCATGCCTCGGTTTTTTCACATTTGGAAGCAGCGCAACAATGT
102001 CTTACATAATATGCGATCGATCACTCTGCAACCAATATTCAAGTACATAG
102051 ACCATGACATCAAAAACATTATCACACCGAGAAGAAAGAAACGTCAATTT
102101 GGTAACCTAATGGCGTTATGCCTGCGGTGAATTCTCCTAAGAGTTCTCCC
102151 AAATTTTATTGATTCCTTGTTTTTAACTTTTTTCGCCAAAGAATCATACAT
102201 ATAGATTTGACACCATTTCAACTTATCAAATACAAGTGAATAAATAATTT
102251 CAAGCTTGAAAGGAATTTAATCATGATCTAAACCTAAACGACAAATTCTT
102301 CACAAGTGAGAATCACTAATTGACTACCCCTTGGTCGCATATACATCAT
102351 GTTGTAATCTGAAAATTGGTTTGGATTTGATCTGATATGTCATTCATAT
102401 AAAACTTGTATTATTTATTTTGAATTTTGCCGCAACAGATAAATCATC
102451 ATCTATTTAGAAAATTTTCATTTGCACCACAATTAATCAGGGGAAAAGGT

FIG. 3 (Continued)

102551 ACAAATTTTCAAATACCACTTATGAGAAGCACTAAGATCACCTTTTTTCTT
102601 TATGACTTTTCTTTCTAAAGCTAAGCTGGTAGTCATGACTCATGATTATCC
102651 TTTTCCTAATGGGAATATTGTGGAAGCGGTTTCAAATCTTTAGACAAAAT
102701 TCCATGGCCACTAAAAGTTAGCAAAGTTAAAATAAGTTTAAAAAATATG
102751 AGTGTACTTGCCCATATGCCATATTGTTGAGATCATAACAAGAGAAATAA
102801 TAGTTTATTGAAGTTTAGATCATAATCACAATACATCATTGCCTTCATCA
102851 ACATTTTCCATGGATTTGAGAGGATCAACTTCAATACTAATGGTGGGGTC
102901 TTATTCATCCATTGCTCTCTAGCCAATTAAGCAGTTAGGTTATTTGTGTA
102951 CTCTAGTAGTTGCCAAATCAATCTTAATATTACAAATGTTGTAATTTCTA
103001 ATTACGTATAGATAAATGACTAGATAAACACGTGGCTTTGGTTTTATCAGG
103051 AAAGTTTTTCAAATCATATATATGAATGTAGAATAGTGTTCTTCATTAAT
103101 TATTAATTAGCATCTCACCATCTGAGACTGGGAGCATGTGACAAGTTGAC
103151 ATGTGTATTAAGAGAACTTTGAGAAAACCACTTTTATGATACTCCCATCT
103201 GAGACTGGGATGAGTACCATTTTATAAAAATATGAGTAGTGAAAAAATAT
103251 TCAAAAAAATTTCTAACATGTCCTTTTAAACATTTTAACTTATAATTTT
103301 AACAAACATCTTCCAATATGCGTTATGAAAACCTTTATAAACTTTTTTAT
103351 AACATGCTTTTGAAAATTTTATAAATCTGTATTTTTTAGAAACAAAGTGAT
103401 ACTTTTGAAAATAGACAAATGAAGTGCTATTTTTTTAAATTTGATATCATA
103451 AGTCTTAACTGTGGTTTGTGTTGAATTTTATTTATATACTTGTCAAATAA
103501 AACTAAATAAATAAATTAAATTATTTTATAATCATGAAGATAATATTATC
103551 ATAAAAGATAAATATAAAATCAACAAATTTATATTTGTTAATAAAAATAC
103601 TTTGAGCTCTTCTTCATAAGACTTTTCCAGCTTCCATCTAGAAAATCACA
103651 TAAATTAAAAGATAAATAACCGAATAAACATAGTTCACATTCTAACTCTT
103701 AGTCTTAGATTTGTTTTAATTTTCAAAGGTTTAGGTATTGTATATGTTTT
103751 TTTTATTGGGTTGCTAGATTTTGATCCAAGAAGAAATGACGGGTTGTAGT
103801 ATAGATGGTTTGTGTTGAGTTTTTTCCCCTTGTTTACTTCGTTTGGTTTT
103851 TGTCCCCAGAATTGTTCTTGTA CTGCTGGTTTATGTCTCTACAAAGTCC
103901 ACGACCATTGCCGGCTCTTGTATTTCAACTTGAATTCTAAATTCGATTG
103951 ATGAAAAAAAATGTATCTCTTAAAGTCCATTAGTACCAAAAATAACTAT
104001 ATCATTACTACATAAAATAGTCTTGGGTTTTTCAAAGTATTTTCGTTGATA
104051 TATGTTAAGAGTTTCGAAATAGACACATAGATATAATGTTGAAATG
104101 TCTCACATAATTATCTCCTTTTCTCTTCATTTCTCTACCTCTCAAGTTTC
104151 CAATCCCACCCTAAGGTAATTTATTTCTTAACCTAAGTAAATTTGTTAAC
104201 AAATCTTAACTAGCTACAAATGTGTATTACAAGTCTTAAATAAAAACCTA
104251 CTTTAATTCAAAGGTATTAAACCTTCCTAAATTGATACTTACTTAGTATC
104301 GATCGGTCTAGTTTAGGGTTTGGACAACACACCATCATGGGGACGAAATT
104351 AGTCATTCTACGGTGTCCAAGACACAAATCTCGGACTCGATGTGGATATG
104401 AACTTCATTATAACTTTTAACTTCATAAAAATACTATTAGGAGGAAG
104451 AATCGGAATCTGCATATCAATCACAATAGACTATAGTATACTTAGATTTT
104501 GATCTAATCAATGGGCTCCTTCAACTAATAAGTAGCCCACTACCAATAAT
104551 GAAATCATAAGACATTATTAATTAATCAATGTTCTAAAAATACTTTGGT
104601 TATGTGTCCCGTAGAGCTAATGTGCACACACAATGAAAGTTGACCCGTTT
104651 CACTTGTCCCACTTTTATGATCTTTTCTTTTAGGTTAAATCCAACCTTTTA
104701 TAATCTCATCTTGTTATCAAACAAAACCTTTTGGCCTGTCTTTTTTCATAAT
104751 TTAAAGTAACTCTCACGGAGAAAAGCCAACATTTTCTTCTTGTTTTATTC
104801 TTTTAAAGAAAAATGAATTCAAGGGGACCCCAAATTTAAAGGAAAACCA
104851 AAACCTCCTTTCTATGTATTTATTACTTGAAGTTTTCTATGTAATCAACAA
104901 TCCTAACAGTAGAGAATAAAAAACATCGTTTTTGGGAGGTTTTATATTAGC

FIG. 3 (Continued)

105001 CCTCTGTCAATGGAGCTATATCACTTGTCAATTTTGCTTAACCCCTTTGCGG
105051 GAAGATTGTTATGAAACAGTTTTAATGGAATTCTAGTTGCCAATGTCACG
105101 TTTAATATGTTTTGTCCCTATACTTTATTGAATCTTATAATCTTTGTTAT
105151 AGAATTATCTACTTTTTAGTATTTTACATTAACATAATCTATAGAATTCTT
105201 CTTTGTCTCTATACAATTAAACAAGTAATATATTCTTAATACATATTAAAA
105251 ATGGTGGTGTGCTATCTGAGCTGTAATAGTTGATTGCTCCAGAGAAGAA
105301 TAGACAAAAATCCTTACTTAAGAGGCCCACTCTGAAAATTTAGACAA
105351 GAAAAATTAAACAAAATTAGGTTACACATATTATCATTTTATATATATGCA
105401 CAACACAAAGTTGACCTTGCAATGTACTATTGAATAAAATAAAATGC
105451 AAGAAGAGAGGGAATTATCACTGTTACCAAGAAAACAACCTCCTCTAAAC
105501 AGGTCTCTATATATATAAACTTTAACACCTAAAGAATTAACACAGATCAA
105551 GAAAAATCCTCAAAACAAAAGTTAAAGCAGAC **ATG** AAG CAA CAG CAA
1 M K Q Q Q
105599 CGT TAC TTG GTC GTC TTC ATC GTC CTT TTA AGC TTT CTT
6 R Y L V V F I V L L S F L
105638 CTG GTAAAGCTTCTTCCTTAATTATATTAACCCCTAATTAAGATCTCATATA
19 L
105691 TCTGAATGTTGTATATATTTGTTGGTATAG TTT GTG AAT CTG AGT
20 F V N L S
105736 GAA GGA AGA ACA GGA GGA GTT GCA GAA GAA TAT TGG AAG
25 E G R T G G V A E E Y W K
105775 AAG ATG ATG AAG AAT GAA CCG TTG CCT GAA CCA ATC AAA
38 K M M K N E P L P E P I K
105814 GAG CTT CTC AAC AAT CCT TTT AGG ACC GCA CAA GAG AGA
51 E L L N N P F R T A Q E R
105853 TTC ATC CAG AAT TTC GAC ACC AAA TCT GTT GTC ATC ATC
64 F I Q N F D T K S V V I I
105892 TAC CAC AAT CCT AAT GAA TAA TCAATGAAGTCTCTCATATAG
77 Y H N P N E •
105934 ATATCTATGACTTTAATTTGTGTTTATGTATGGATCGACTTATACGTGCA
105984 CGTATATGTTATTAATTAAGAAAAGAAAAGCTGCTTGAGTTGTTGTGTT
106034 ATACACGTATACTAAATATGTTCTGTTTAGTGCAGAAATGTTAACCCTAG
106094 CTATAAGGGATTTTTTGTCTTTTTTTTTTTGTTACCATTAATGTGAGTGA
106144 GTGAGTTTTGTGTGATGAAAATTAGATTTGCTTCACATTTTGTTTTGATA
106194 TATATAAATCAATATACTGTGCCTTTCGTGTCTTGTTTCTTATATTATTT
106244 TGTGACATTAATTAATTATCTTATCAAAAATTTATTTTATTAAGTGTGTC
106294 CTATGGAAAAAGATGAACAATATGAGTTAACCTCATCTCAAGGAGATTCT
106344 TTTTGTGTTTGTGTTTTC

FIG. 4

1 AAGCTTTACAAATGTCCCCCAAAGATGAAACCACGTTATTATTAGTAAATCCTGAAAAGG
61 TTAACGCTTCTGTTCCCTCGAATTCTAAACCATCTGAAATATCTAGTGGTTTAAAATGGAG
121 ACTTGAGGATATAGTCTCCTGAACCAGCTGTCACGGCTGAGTTAGATAACATTACTGAAT
181 TTCTACGGGAGCGGTTGAAATCACTTTCGCCCTTTAAGAAGAAGCCTACACCGGGCACC
241 TTCTTTACGCAATTTCGAAATTTAGTCTTGCCAGGCAGTCGTTGGATCGAAGGTCTTTTTC
301 GATACCGAGGAATCTGACTTTGCAAGGAATAATTCCTAATCACACCACCCCAACCCCTGA
361 ATACACTTCAGGACCCTCTGAAACCAACTTCGTTTTCGGCTAAATCACAAGAATCTCCAC
421 TCATTCCGATTTTAGCCAATTAATATGATATCGGTCTGGGAAGCCGATAAGGAAATTCT
481 ACAAAAAGAGTTTATGAATGAGGAAAATAAGGAAAAGAGAGAACTATTTTTTAGGTACCC
541 TGAAAGAGAACGAGAAAAATTTAGAAAAAATACTACTCTCATCTGTACACTGTTCAAAA
601 GAATATCCnnnnnAATGGTTAGATAATATAAGAAAAGGATAAGTATGATTAAACTGAAAC
661 CACGTCGGCAGAAACAAAGTGAATTCCCCCTTTAGAGGAAGTTCGTTTCTTAAATATAG
721 AAAACAAAGAAGTAGTCGCCTCCCTTTTAAATGATCTCAGAAAAACGAGAAGTAAGTA
781 TAAAAGATATTCAAAATCTACACAGTCAACTAAATTTTACTAATCAAATGCTTTTTCAAT
841 TAGCAAATAAAAAACAAAAGAAAAAAGmGAAATTTGAAGAAAAATCGTTAATAAAACCAT
901 TTAAATTCTCAGAAGAAGAGATAAAACAGTTAAAAATTGGTCAAACTTTGGATTCTTTAT
961 ACGATGAAGTAAAACAAAAGTTATCTATCTCGGTAAATAAAAGAAAAACCGAAATCTAATA
1021 ATGATATGCCCAAAGGACAAATCCAAATCAAGAAGTTTTAGACGAAATCGAAAAGAGAT
1081 TAAAACAAACTCTGAACGACACAATAAATGTGATAGAAGAACTAAAACTCAGACTCAT
1141 GTTCAGAGTCTCCCGATCGTATTGAAAAATAAAACGTAATAAATCAGAGATTTCCAGTA
1201 AGCCGAAATTTTTACACTCGCCCCACCTTCGATATCATCGAGATGGCGATGGACACCTCA
1261 GCATTGATGGAATGGATACTGAGTGATATGATGGATGACAGATGATGAATATAGAAAAAC
1321 TCACGAAATAACAATGGCCGCTACAGCATATAGAGTAAACATACCGAGGAACAAACAAT
1381 AAAATTAATTATATCTGGATTACGGGAGTATTAAAAGGCTGGTGGGATAATTACCTCAT
1441 GCCAGAACAAAAGAATTATGTTCTAAGCTGTGTAAAAATAGAAAACGAAGAAGGAATACC
1501 ACTAATGGTGGAAACATTGGTGGTAGCAATAATTCATAACTTTATAGGAGATCCAAAGAT
1561 TTTTGAAGAAAGAACATCTTTATTACTTCATAATCTAAGATGTCCAACCTTAGGTGACTT
1621 TAGATGGTATTTCAGAAAATTTTTTAGCTATGGTTTTTAAACAAGGGAAGATTGTAGAGAACC
1681 TTTCTGGAAAGAACGGTTTTATAGCTGGATTACCGGATATCTTTGCTGAAAAGGTAAAAGA
1741 AAATTTACAAAAGGAATGCCCAAACACCCAATTAAAAGATGTACCATACGGGAAAATAAG
1801 TTCAGTTGTAAAAAATACAGGTCTTCAGTTATGCAATAATATGAAAATAGAAAATAAGAT
1861 AAAAAAGAGTGAGAGTCAGGGCATCAAGGAATTAGGGGAATTTTGTACTCAATACGGTTA
1921 TGAACGAAATACCCCTCCATCAAAAAATAAAAAGAAAAATAGCAAAAAGAAGAACAgGGAG
1981 AAACAAGCGCTAAAACAAGCGCTAAACCAGCACGTAAAAATTTTAGAAAAACGGTTAATT
2041 TTAGAAAACCATGAAAGTCTAATGATAAGCCCACTATAGTCTGTTATAAATGTGGACGCA
2101 TAGGACACATGAAGCGAGACTGTAGACTAAAAGAAAAAATTAGTAATTTGACCATAAGTG
2161 ATGAATTAAAAGAACAAATGGAAAAACTTCTGATAAATTCCTCCAGAAGAGGAAGAAACA
2221 GAAGAATCAATAGGAGATTCTGATTACGAAGTATTGGACATGAGGATAACAATTGTAATT
2281 GTGTCTATAAAATAAATACGATAAGTAGTGAATTAATAATTTGCGTTAGATTGCATTGATA
2341 AAATTAATAATCCGGAGGAAAAGACCAAAGCCTTAATAGACATGAAAAGGCTACTCGTTG
2401 AAAAAGATGAACCCAGTTCATCTTCACAAAAACCTGAATTTATAGGATATGATTTTAAAG
2461 AAATATTGAGAAAAGCGAAAACATCACATAAAGAAATAACCATTAGCGATCTTAATAGTG
2521 AAATAAATAAATTAAGCCGAAATCGAATCTATAAAAGTCGAGCTACAAGAATTAAAG
2581 ATAAAATTATACATGAGGAATCCATCTCCTCTGCCGACGAAAATTCACAAGAAGAGGAAG
2641 CTAGTAGACCTTCCATCAAAGAAATAACATACAAAAGACAAAAGTGGCATGTAAAAATAG
2701 CCCTAGAATTTGTTTGTGTTTGTGACCGTTTCATTGTGGTCAAAGATGAGTCCTTACCTAA
2761 CACAATAAAAAACGTTACTCTTAAATATCAAAGGAGAGCTACAAATATCAATGAATGAAT
2821 GACATTAATATTTTTCTTTAGTTTTTAAACTTGAATGAGTTGTTTTCATAAATATCTGAC

FIG. 4 (Continued)

2941 ATATTTTTATCTCAAAGTCAACGAAATATTATAAAAGAATCAATTAAAAAAATTATTCT
3001 TTTGCAGAAAAAAATTAATAATATGAACTCCTCCACACCATATTACCATATTATAAA
3061 TATAAAAAAACCTCTCACAAATGTGCATTCTGGAATTCTTTATGTTGAGAGATTAATCTC
3121 TAAAGAAAAAAGGTTGAGAAAGGTGCAGCAACA **ATG** TCT CCA TTC TGT AGA
1 M S P F C R
3172 AAC TTT TCA ATG GCA TGG GTG CTT ATG GCA TTT GTG TTG TTT
7 N F S M A W V L M A F V L F
3214 GCA AAC AGT GCT ATG CCC ACA AAT GGA TCC ACT GTT GGG GTA
21 A N S A M P T N G S T V G V
3256 AAA AAC ATG TTG GGT GGT AAA TTG ATG CTA AAC GTT TTA TGT
35 K N M L G G K L M L N V L C
3298 CCC CAT ATT GAT AAG CAA CAC ATT ATC CCG AAT GGT GGT TCA
49 P H I D K Q H I I P N G G S
3340 TTT GAG TGG AAG TAC AAT GGT GGT GCT CCA CCA ATA GGA CAA
63 F E W K Y N G G A P P I G Q
3382 TCA CCA TTC ATG TGT TTC TTT CGG TGG AAT AAT GTT CAT CAC
77 S P F M C F F R W N N V H H
3424 TCC CTT GAT CTG TGT TCA CCA AGC AAG TAT ACT GGT TGT GAA
91 S L D L C S P S K Y T G C E
3466 AAT GCC ATT TGG GAA ATC AAA GAA AAG CAA TTT TGT AGG TAC
105 N A I W E I K E K Q F C R Y
3508 AGA GGT GGA CCT ATT AAT TAT TTT TGC TAT GAC TGG GAT GAT
119 R G G P I N Y F C Y D W D D
3550 TAG TTATATAGATTATTCATGTTTCATCTCAATAAAAAAATGACTTTAGAGTGATTCTT
3609 AGTTTGCTTAACATTCTTACATATTCCTAACTATTCCGTCACTACCACCCGTAACCTATAT
3669 TTATTTTAAATTAGTATCTGTCTCACAGTTTTATTTTTTAAAAAAGGTTATGTGGATTAGAAG
3729 AGAGATAAATATGTAGACGGTCACCAACCTTAATTTTTTGAAGTATGTAAGACTATATTGA
3789 CCAAGAATATATGTTTAACTCATTCATTTAAAGACTATATCTCCATTTATGATTATGCA
3849 AATGCAATTAGTTTTTTTTTTTCATTGAAGAATTCAAAGAAAGTTATCATTAAAAAGTAT
3909 CATTAAATCACTTATATGTTGTTTCTTAATATCCTTATTGTTAATAGAATAATTTTTTTTT
3969 ATCCTTTAATTAAGGTTATTACTACTTTTTTTTTTCATATCTTCATTATTTTGAAATATTTT
4029 TAAAATTTATCAATTTTTTGTAACACCCCAGAAAATACATGTAAGTATCACTTTTTTTTTTA
4089 TATTACAAATTTATGACTTATAGAAATACAAATATTAATAAATATAAGGTTCAAACTACA
4149 TCCTAAAGTCTTTCAGACCCCTCTGACACATGTATCATCTGCTCGTATATGTGATACAGTC
4209 ATCGCAGTTCACAAGATAACAAGAAAACCAAGGGTAAGCTAATGAAAAAAATTCCATAA
4269 CATATTTAATTCATGCAAAAAGAACCAGTCAAAGTAATCATTTTATAAACATTTCTTTAAA
4329 TATTGTTATATAAAATTTCAATATCAATTTTCATCATTCATATAGACCACACATGGATCTA
4389 TTTTCAATCACAATCATTGGATTTTCATTTTAATCCTACTTCGnCTTCCAGAAGACTCATT
4449 AAGTATGCCCTACCAGAGACTAACACCTAATCAAAGAGAAATGATCAAGGTAAGTTCAA
4509 ACATCCAATAACGAGTGCCTACAGTGGGACCCAATGTGTATGAACTCCTTATCAGCTTCT
4569 CACCACCTGATATCTTATTCTATATGACGTAGATCATCAGTGAACTAGAGGATCTCCGT
4629 TAAACATATGTTTTTTTATACTTAATGTCATCAAACAACACTCACACATTATCCCAAATG
4689 TATGACATCAATTTTCATACAATTTTCATCATTCATATATAATACATATCATTGAATCACA
4749 TAACATTTTAAAAATTCATACCATTCAAGAACTTTTCCAACATCAAAGCAATATTTACTT
4809 TCAAACCTATCAAAATATAATTATTATTTAATAAAGCTt

FIG. 5

142000 TTATCTTATTTCCATATAAATTGTTGTTTTACTTTTCAAAATTTTTTAATTTT
141950 TTATATTTATCTTTTTTACAGTTTAAAATTAATAAAATGAAACTTTTTTTTC
141900 TTAAATGTGTTAAATATAAAATCAAAAAAGTTGTTATATGGTACATGGC
141850 ACAATCTTATAAATTATTAATTTGAAAACGATACTTTATATAATAAAATT
141800 ATCTTAGTTGACATTTTTTATTAGTGTTTTCAATCATATTTTTGTTTGCTT
141750 GATAAGCGTAAAACAAATCAAACCTAACGATACTTTATATAATAAAATTA
141700 TCTTAGTTGACATTTTTTATTAGTGCTTCAATCATATCTTTGTTTGCTTG
141650 ATAAGCGTAAAACAAATCAAGTAAAGTTGGGCACCTCAATTGTTTTAAAA
141600 AAGTTTGGGTACCTCAAAAATTAATAGGTCTTGTCAGATTCTTACAAAAA
141550 AAATCTGGAAGAATTTATGAAAGAAGGGGGGGGAGGGGGGGGAGGGGGGGG
141500 AAGTGAAGATGAATATTCAACAAAAGAGGGTAGGCATGATGTTAAGTGAG
141450 TTAAAAAACTATGTTAATGGAGACAATTTTCTGTTAACAAACCCGTTAAT
141400 TGAAAACGATAGCATTTCTTCTCTAACAAATGTAAAACGATATTGTTTTATC
141350 ATAACACTCATTAATTTTCTGAGTTTCAAATCATATAAAGATTTAGGGG
141300 GGTGTATTCAATTAAGGATTTGAAATGATTTGTATTAAATGACAAATCC
141250 CATGTTATTTCAAACATGAATTGTAAAAACTTTTTTAAATCAAGTGTTA
141200 TTAGATTAGTGATTTTAAATGTACAACCAACCCACTGTTATTGGAAAC
141150 ATTTTAAGTAGTGGATTTAAATGACTTGAGTGATTTTGGGTGGGATTGC
141100 AGAAAATTTCTTAGTTAAGAATTCAAACATCCAAATCTCATGGTTTCAAG
141050 TAGAATTTGGGAGAATTTTAATAACAAATCTCCTAATTTACCAAAAGTCA
141000 CCAAATCATTTAAAACTCATTAATTTTAAATGATTTCAAATCTCCAG
140950 TTGAATACATCCCCTTGAATTAGAGATTTTGCTCGATTTGGGACCTAAG
140900 ATTGAATTTTGGGGATTTAGTTTAATCGTTACAACAAAATGACATCGTAT
140850 TATTGTTATAGGAAACAATGTCGTTTTTCAGTTGACATGTATGTTAATAGA
140800 AAATTAACCTCTATTAACGGGATTTGCTAACCCATTTAACATCGTAACTAA
140750 ATGGTCAAGTCAATAAAAGTTTGGTATTTATTTGAAAAGTCAACGTAAGT
140700 TTGATATTTATTTGAAAAGTCAACATAAATTTGATATCTTATTTGTTTTT
140650 GACAGACATAAGGATTTACATCAATGTTTTTAATAAATTAAAGATTATTA
140600 TGACATTTTTTCCATTTAAAATTGCCAATGTTTTTCGAAACCAAGATACTC
140550 AAAATTGACATACCTAATTCAATCTACATTTGTTTGACAGCAATTCACGT
140500 GCCTTGACCACATGGCACATACTGGCAATACATCAATTTTAAGGAAAAGG
140450 TAGATTTCGGATACAATATAATGGAAATAAGTGGAAAGGATCATTGACTAC
140400 TTGACTTGTAACAAACAACACACAGTATATAACTCATTCGACATTTACAA
140350 ACAACATTGTGCTAGCTTAAACTCCCTCTCCTATTCAAAAAA **ATG**

	1											M
140305	GAT	ATT	CCA	AAG	CAA	TAT	CTA	TCA	CTA	TTC	ATA	TTG
2	D	I	P	K	Q	Y	L	S	L	F	I	L
140269	ATT	ATC	TTC	ATA	ACT	ACA	AAA	TTA	TCA	CAA	GCC	GAC
14	I	I	F	I	T	T	K	L	S	Q	A	D
140233	CAT	AAA	AAC	GAC	ATT	CCA	GTT	CCC	AAC	GAT	CCA	TCA
26	H	K	N	D	I	P	V	P	N	D	P	S
140197	TCA	ACA	AAT	TCT	GTG	TTT	CCT	ACC	TCG	AAA	AGA	ACC
38	S	T	N	S	V	F	P	T	S	K	R	T
140161	GTG	GAA	ATC	AAT	AAT	GAT	CTC	GGT	AAT	CAG	CTA	ACG
50	V	E	I	N	N	D	L	G	N	Q	L	T
140125	TTA	CTG	TAT	CAT	TGT	AAA	TCA	AAA	GAC	GAT	GAT	TTA
62	L	L	Y	H	C	K	S	K	D	D	D	L
140089	GGT	AAC	CGG	ACT	CTG	CAA	CCA	GGT	GAG	TCG	TGG	TCT

FIG. 5 (Continued)

140053 TTT AGT TTC GGG CGT CAA TTC TTT GGA AGG ACG TTG
86 F S F G R Q F F G R T L
140017 TAT TTT TGT AGT TTT AGT TGG CCA AAT GAA TCG CAT
98 Y F C S F S W P N E S H
139981 TCG TTC GAT ATA TAT AAA GAC CAT CGA GAT AGC GGC
110 S F D I Y K D H R D S G
139945 GGT GAT AAC AAG TGC GAG AGC GAC AGG TGT GTG TGG
122 G D N K C E S D R C V W
139909 AAG ATA AGA AGA AAC GGA CCT TGT AGG TTT AAC GAT
134 K I R R N G P C R F N D
139873 GAA ACG AAG CAG TTT GAT CTT TGT TAT CCT TGG AAT
146 E T K Q F D L C Y P W N
139837 AAA TCT TTG TAT TGA CAACAATATGCTGATGTTCTGTCTTTTAC
158 K S L Y •
139793 GACTCATGGAGTTTCATTGTTTGAAACAATAATATAAAACATATAAAATT
139743 TCTATTATTCCAAGTTCCAAGTTATAATAATTTGATAATCATATCATATT
139693 ATCATCTTAAGCATTCAATGCTACAAAGATAATACCCCCAAGCTATTTTA
139643 CATTAAAAGCTGAAACAGAGACACAATACTAACGATAAAAAGTTCGT
139593 TCTTTATGCAACCATACATACATATACACAAAGATAGACAGGTAGTGTCC
139543 TAATAATTCTACTTGGGTGAGGTATGAACAGCAGCAACAGTAGATACCA
139493 TGTATCCATACCACACATATTATGAGGCCCTCTGCAGATTTTGTAGTAAC
139443 CATGCTCTCCCCACATCGCTCCCCACGAGTTCTTGATAATCCAA

FIG. 6

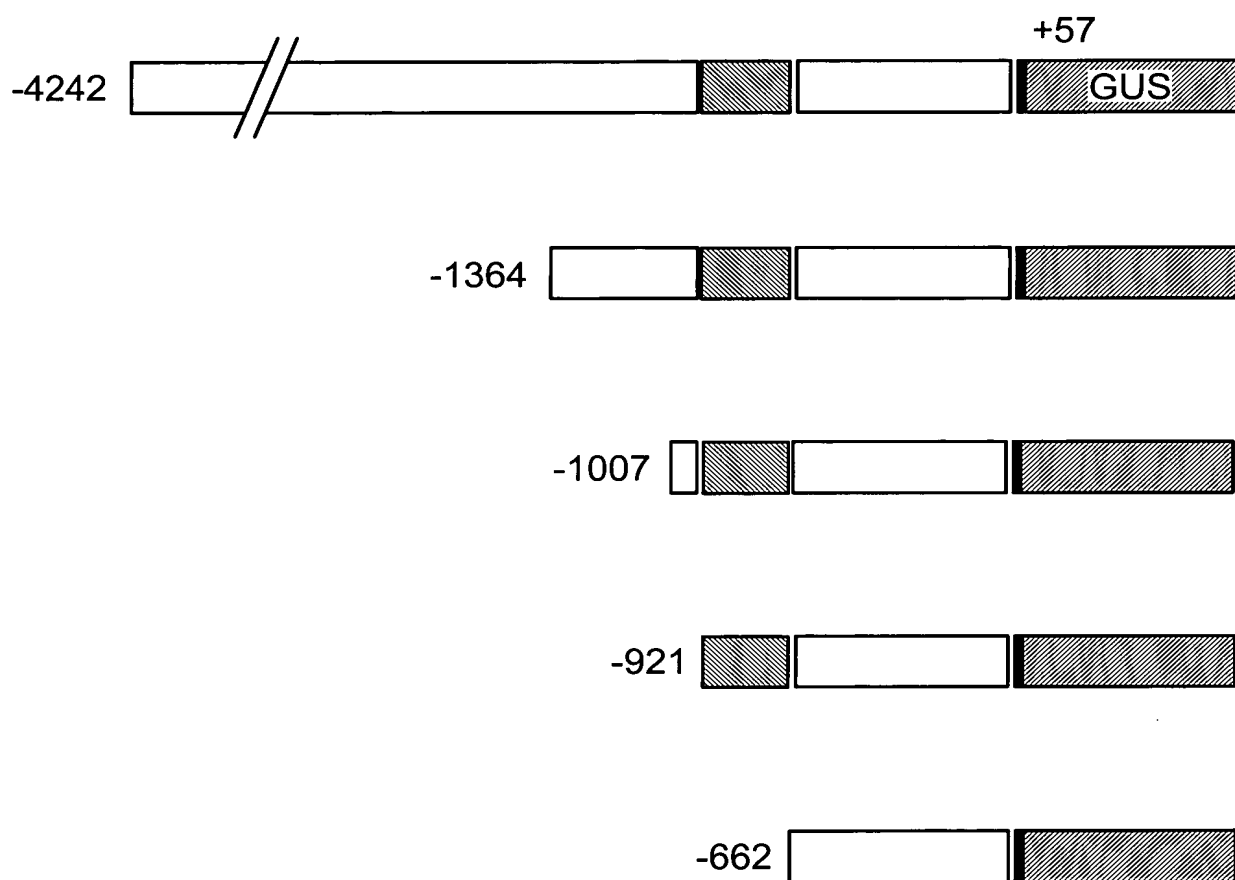
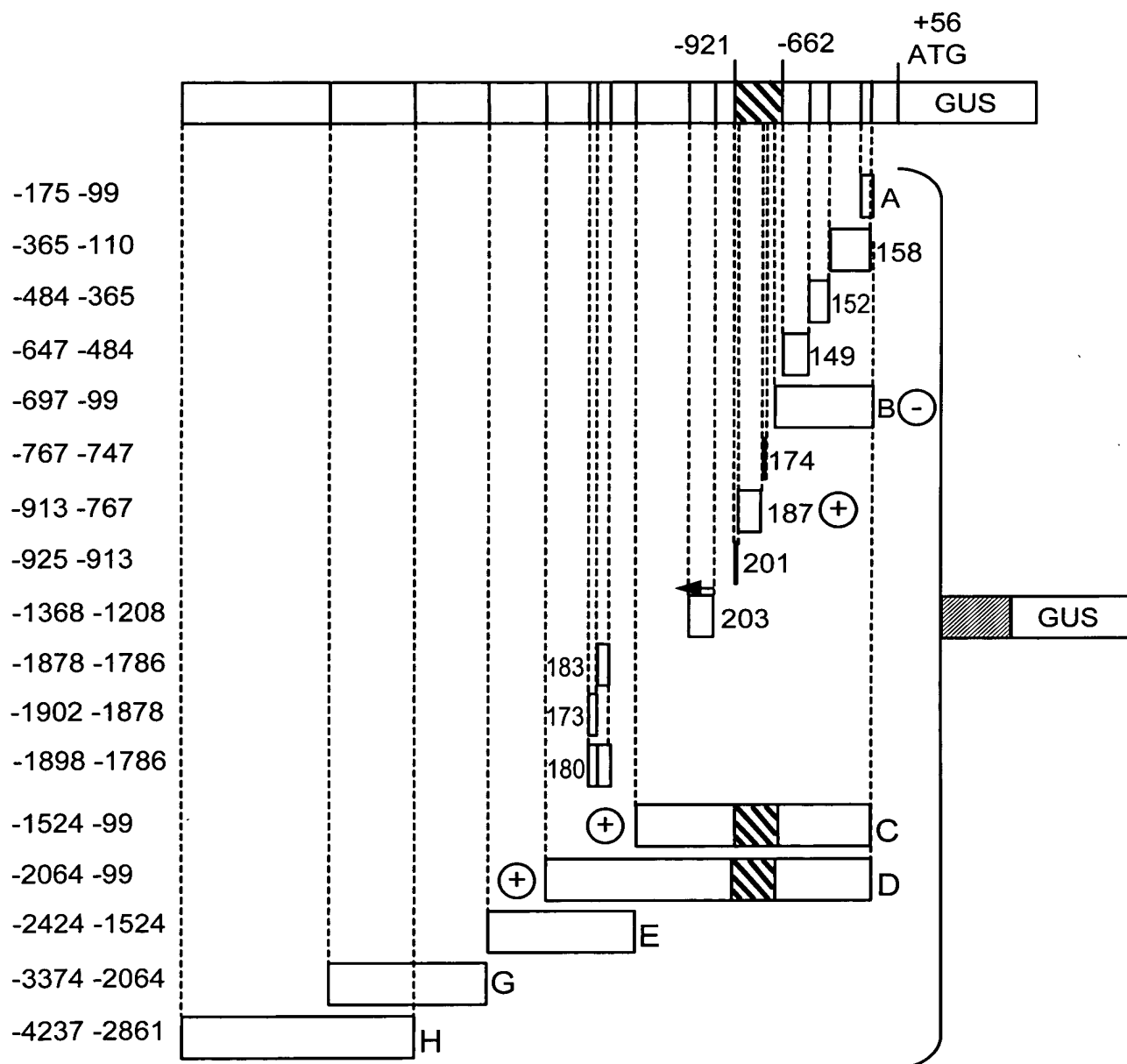


FIG. 7

G564 promoter:
Gain of function constructs



(+) = Gus Activity in Suspensor

(-) = No Gus Activity in suspensor

▨ = Region Required for Gus Activity in the Suspensor

Web Signal Scan Program

Database searched: PLACE

URL: <http://www.dna.affrc.go.jp/btdocs/PLACE/>

This is the sequence you submitted

>G564 promoter (-921 to -662), 450 bases, 3D1A0BF4 checksum.

TGAAAAGTGAAGAAAACCATGTAATGAAAACAAAATGGCAGACAATCAA
 AAAAAGTTTTTCACGCAAAAATTTTCTTCAAAATTTATAACATTTTCATGTT
 GTGTTTGTTCAAAGCCTAGAAAAACGAAGAGTTACTATTGGTAATGAAA
 AGCGAAGAAAACACATAATAAAAACAAAATGGCAGACAATCAAGAAAA
 AGTTTTTCACACAAAACCTTTTTTCAAAATTTACTATGTTTATTTTCGAAATT
 TAGAAAAACGAAGAGTTATTATTAGTAATGAAAAGCGAAGAAAACACTACGT
 AATAAAAAACAAAATGGCAGACAATAAAAAAGTTTTTCACGCAAAAATTT
 TCTTGGTGCGCAGAAAGTTATATATATTAATTAATTAATTTTCATTTACT
 TTTTTCCTTTTATTTTAAAGTTAAATTATTATTATTTTCATTTAAAAT

Notation: H = A, C, or T

R=A or G

K = G or T

W=A or T

RESULTS OF YOUR SIGNAL SCAN SEARCH REQUEST

/tmp/signalseqdone.9437: 450 base pairs

Signal Database File:

Factor or Site Name	Loc.(Str.)	Signal	Sequence	SITE #
-300ELEMENT	site 1	(+)	TGHAAARK	S000122
2SSEEDPROTBANAP	site 101	(-)	CAAACAC	S000143
ACGTABOX	site 296	(+)	TACGTA	S000130
ACGTABOX	site 296	(-)	TACGTA	S000130
AP3SV40	site 159	(-)	TGTGGWWW	S000169
CAATBOXI	site 44	(+)	CART	S000028
CAATBOXI	site 189	(+)	CHAT	S000028
CAATBOXI	site 323	(+)	CAAT	S000028
CAATBOXI	site 138	(-)	CHAT	S000028
CANBNNAPA	site 101	(-)	CNAACAC	S000148
CCAATBOXI	site 138	(-)	CCAAT	S000030
CEREGLUBOX2PSLE	site 55	(-)	TGAAAAC	S000033
CEREGLUBOX2PSLE	site 201	(-)	TGAAAAC	S000033
CEREGLUBOX2PSLE	site 333	(-)	TGAAAAC	S000033
DOFCOREZM	site 4	(+)	AAAG	S000265
DOFCOREZM	site 53	(+)	AAAG	S000265
DOFCOREZM	site 112	(+)	AAAG	S000265
DOFCOREZM	site 149	(+)	AAAG	S000265
DOFCOREZM	site 199	(+)	AAAG	S000265
DOFCOREZM	site 282	(+)	AAAG	S000265
DOFCOREZM	site 331	(+)	AAAG	S000265
DOFCOREZM	site 364	(+)	AAAG	S000265
DOFCOREZM	site 419	(+)	AAAG	S000265
DOFCOREZM	site 216	(-)	AAAG	S000265
DOFCOREZM	site 399	(-)	AAAG	S000265
DOFCOREZM	site 408	(-)	AAAG	S000265
GT1CONSENSUS	site 120	(+)	GRWAAW	S000198
GT1CONSENSUS	site 141	(+)	GRWAAW	S000198
GT1CONSENSUS	site 196	(+)	GRWAAW	S000198
GT1CONSENSUS	site 253	(+)	GRWAAW	S000198

Figure 8

GT1CONSENSUS	site	69	(-)	GRWAAW	S000198
GT1CONSENSUS	site	90	(-)	GRWAAW	S000198
GT1CONSENSUS	site	347	(-)	GRWAAW	S000198
GT1CONSENSUS	site	388	(-)	GRWAAW	S000198
GT1CONSENSUS	site	436	(-)	GRWAAW	S000198
GT1CONSENSUS	site	218	(-)	GRWAAW	S000198
GT1CONSENSUS	site	401	(-)	GRWAAW	S000198
GT1CONSENSUS	site	402	(-)	GRWAAW	S000198
MAMMALENHAN	site	158	(-)	GTGGTTTK	S000121
MARTBOX	site	324	(-)	TTWTWTTWTT	S000067
MRE1	site	356	(-)	TGCRCNC	S000068
NTBBF1ARROLB	site	418	(-)	ACTTTA	S000273
POLASIG1	site	168	(+)	AATAAA	S000080
POLASIG1	site	301	(+)	AATAAA	S000080
POLASIG1	site	324	(+)	AATAAA	S000080
POLASIG1	site	237	(-)	AATAAA	S000080
POLASIG1	site	411	(-)	AATAAA	S000080
POLASIG3	site	268	(-)	AATAAT	S000088
POLASIG3	site	427	(-)	AATAAT	S000088
POLASIG3	site	430	(-)	AATAAT	S000088
POLASIG3	site	433	(-)	AATAAT	S000088
POLLEN1LELAT52	site	11	(+)	AGAAA	S000245
POLLEN1LELAT52	site	119	(+)	AGAAA	S000245
POLLEN1LELAT52	site	156	(+)	AGAAA	S000245
POLLEN1LELAT52	site	195	(+)	AGAAA	S000245
POLLEN1LELAT52	site	252	(+)	AGAAA	S000245
POLLEN1LELAT52	site	289	(+)	AGAAA	S000245
POLLEN1LELAT52	site	362	(+)	AGAAA	S000245
POLLEN1LELAT52	site	71	(-)	AGAAA	S000245
POLLEN1LELAT52	site	349	(-)	AGAAA	S000245
PYRIMIDINEBOXHV	site	400	(+)	TTTTTTCC	S000298
RAV1AAT	site	97	(-)	CAACA	S000314
ROOTMOTIFTAPOX1	site	374	(+)	ATATT	S000098
SEF4MOTIFGM7S	site	170	(-)	RTTTTTR	S000103
SP8BFIBSP8BIB	site	134	(+)	TACTATT	S000184
TATABOX2	site	81	(-)	TATAAAT	S000109
TATABOX3	site	375	(+)	TATTAAT	S000110
TATABOX4	site	368	(-)	TATATAA	S000111
TATABOX5	site	238	(+)	TTATTT	S000203
TATABOX5	site	412	(+)	TTATTT	S000203
TATABOX5	site	434	(+)	TTATTT	S000203

For more information about the SignalScan Program, please contact Dr Dan S. Prestridge Tele:(612) 625-3744 Advanced Biosciences Computing Center. E-mail:danp@biosci.umn.edu 1479

Gortner Ave. University of Minnesota St. Paul, MN 55108

The TFD data is at the gopher site, gopher://genome-gopher.stanford.edu. For more information about the WebSignalScan service, please contact Meena Sakharkar, meena@biomed.nus.sg, BioInformatics centre, NUS .

Database Searched: PlantCARE

URL : <http://sphinx.rug.ac.be:8080/PlantCARE/>

Sequence submitted:

>G564 promoter (-921 to -662) 11/21/00

```
+ GAAAAGTGAA GAAAACCATG TAATGAAAAC AAAATGGCAC GACAATCAAA AAAAGTTTTC ACGCAAAATT
+ TTCTTCAAAA TTTATAACAT TTTCATGTTG TGTTTGTTTC AAAGCCTAGA AAAACGAAGA GTTACTATTG
+ GTAATGAAAA GCGAAGAAAA CCACATAATA AAAACAAAAT GGCACGACAA TCAAGAAAAA GTTTTCACAC
+ AAAACTTTTT TCAAAATTTA CTATGTTTAT TTCGAAATTT AGAAAAACGA AGAGTTATTA TTAGTAATGA
+ AAAGCGAAGA AACTACGTA ATAAAAACA AAATGGCACG ACAATAAAAA AAGTTTTTTC GCAAAATTTT
+ CTTGGTGC GC AGAAAGTTAT ATATATTAAT TAATTAATTT TCATTTACTT TTTTCCCTTT TTATTTTAAA
+ GTTAAATTAT TATTATTTTC ATTTAAAA
```

```
- CTTTTCACTT CTTTTGGTAC ATTACTTTTG TTTTACCGTG CTGTTAGTTT TTTTGAAAAG TGCGTTTTAA
- AAGAAGTTTT AAATATTGTA AAAGTACAAC ACAAACAAAG TTTCGGATCT TTTTGCTTCT CAATGATAAC
- CATTACTTTT CGCTTCTTTT GGTGTATTAT TTTTGTTTTA CCGTGCTGTT AGTTCTTTTT CAAAAGTGTG
- TTTTGAAAAA AGTTTTTAAAT GATACAAATA AAGCTTTAAA TCTTTTTGCT TCTCAATAAT AATCATTACT
- TTTGCTTCT TTTGATGCAT TATTTTTTGT TTTACCGTGC TGTTATTTT TTCAAAAGTG CGTTTTTAAA
- GAACCACGCG TCTTTCAATA TATATAATTA ATTAATTAAA AGTAAATGAA AAAAGGGAAA AATAAAATTT
- CAATTTAATA ATAATAAAAG TAAATTTT
```

3-AF1_binding_sit

Site Name	Organism	Position	Strand	Core simil.	Matrix simil	sequence
3-AF1_binding_sit	ST	260	+	1.000	0.860	AAGAgttatt

Function:

AAGAA-motif

Site Name	Organism	Position	Strand	Core simil.	Matrix simil	sequence
AAGAA-motifAvena	sativa	6	+	1.000	0.903	gtgAAGAA
AAGAA-motifAvena	sativa	151	+	1.000	0.870	gcgAAGAA
AAGAA-motifAvena	sativa	284	+	1.000	0.870	gcgAAGAA

Function:

ABRE

Figure 8 (cont'd)

Site Name	Organism	Position	Strand	Core simil.	Matrix simil	sequence
ABRE	Hordeum vulgare	293	+	1.000	0.854	actACGTaat
Function: cis-acting element involved in the abscisic acid responsiveness						

ACE

Site Name	Organism	Position	Strand	Core simil.	Matrix simil	sequence
ACE	Petroselinum crispum	293	+	1.000	0.908	actACGTaat
Function: cis-acting element involved in light responsiveness						

AE-box

Site Name	Organism	Position	Strand	Core simil.	Matrix simil	sequence
AE-box	Arabidopsis thaliana	67	-	1.000	0.852	AGAAaatt
AE-box	Arabidopsis thaliana	345	-	1.000	0.852	AGAAaatt
AE-box	Arabidopsis thaliana	361	+	1.000	0.852	AGAAagtt

Function: part of a module for light response

AT1-motif

Site Name	Organism	Position	Strand	Core simil.	Matrix simil	sequence
AT1-motif	Solanum tuberosum	409	+	1.000	0.859	ttttATTTaaa
Function: part of a light responsive module						

Box_4

Site Name	Organism	Position	Strand	Core simil.	Matrix simil	sequence
Box_4	PC	375	+	1.000	1.000	ATTAat
Box_4	PC	379	+	1.000	1.000	ATTAat
Box_4	PC	383	-	1.000	1.000	ATTAat

Function:

Box_I

Site Name	Organism	Position	Strand	Core simil.	Matrix simil	sequence
Box_I	PS	107	+	1.000	1.000	TTTCaaa
Box_I	PS	203	+	1.000	0.857	TTTCaca
Box_I	PS	219	+	1.000	1.000	TTTCaaa
Box_I	PS	240	+	1.000	0.857	TTTCgaa
Box_I	PS	241	-	1.000	0.857	TTTCgaa
Box_I	PS	249	-	1.000	0.857	TTTCtaa

Figure 8 (cont'd)

Function:

Box_II

Site Name	Organism	Position	Strand	Core simil.	Matrix simil	sequence
Box_II	ST	139	+	1.000	0.889	TGGTaatga
Box_II	AT	161	+	1.000	0.954	CCACat_aat

Function:

CAAT-box

Site Name	Organism	Position	Strand	Core simil.	Matrix simil	sequence
CAAT-box	Hordeum vulgare	43	+	1.000	1.000	CAAT
CAAT-box	Arabidopsis thaliana	137	-	1.000	1.000	aCCAAt
CAAT-box	Hordeum vulgare	188	+	1.000	1.000	CAAT
CAAT-box	Hordeum vulgare	322	+	1.000	1.000	CAAT
CAAT-box	Arabidopsis thaliana	351	-	1.000	0.857	aCCAaG

Function: common cis-acting element in promoter and enhancer regions

ERE

Site Name	Organism	Position	Strand	Core simil.	Matrix simil	sequence
ERE	Dianthus caryophyllus	239	+	1.000	0.875	ATTTcgaa
ERE	Dianthus caryophyllus	241	-	1.000	0.875	ATTTcgaa
ERE	Dianthus caryophyllus	413	+	1.000	0.875	ATTTtaaa
ERE	Dianthus caryophyllus	441	+	1.000	0.875	ATTTaaaa
ERE	Dianthus caryophyllus	442	-	1.000	0.875	ATTTtaaa

Function: ethylene-responsive element

G-box

Site Name	Organism	Position	Strand	Core simil.	Matrix simil	sequence
G-box	Zea mays	17	+	0.842	0.870	CATGta
G-box	Zea mays	38	+	1.000	0.903	CACGac
G-box	Zea mays	94	+	0.842	0.886	CATGtt
G-box	Zea mays	183	+	1.000	0.903	CACGac
G-box	Zea mays	317	+	1.000	0.903	CACGac

Function: cis-acting regulatory element involved in light responsiveness

GC-repeat

Figure 8 (cont'd)

Site Name	Organism	Position	Strand	Core simil.	Matrix simil	sequence
GC-repeat Function: ?	Oryza sativa	351	-	1.000	1.000	gCACCaag

HSE

Site Name	Organism	Position	Strand	Core simil.	Matrix simil	sequence
HSE	Brassica oleracea	49	+	0.944	0.878	aAAAAggtt
HSE	Brassica oleracea	50	+	0.944	0.912	aAAAAgttt
HSE	Brassica oleracea	52	-	0.944	0.874	gAAAActtt
HSE	Brassica oleracea	66	-	1.000	0.978	aGAAAattd
HSE	Brassica oleracea	77	-	0.833	0.868	aTAAAtttt
HSE	Brassica oleracea	87	-	1.000	0.853	tGAAAatgt
HSE	Brassica oleracea	196	+	0.944	0.912	aAAAAgttt
HSE	Brassica oleracea	198	-	0.944	0.874	gAAAActtt
HSE	Brassica oleracea	210	+	0.944	0.874	cAAAActtt
HSE	Brassica oleracea	212	-	0.944	0.912	aAAAAgttt
HSE	Brassica oleracea	213	-	0.944	0.878	aAAAAagtt
HSE	Brassica oleracea	327	+	0.944	0.878	aAAAAagtt
HSE	Brassica oleracea	328	+	0.944	0.912	aAAAAgttt
HSE	Brassica oleracea	330	-	0.944	0.874	gAAAActtt
HSE	Brassica oleracea	344	-	1.000	0.978	aGAAAattd
HSE	Brassica oleracea	361	+	1.000	0.888	aGAAAgtta

Figure 8 (cont'd)

HSE Brassica 385 - 1.000 0.853 tGAAAatta
oleracea

Function: cis-acting element involved in heat stress responsiveness

I-box

Site Name	Organism	Position	Strand	Core simil.	Matrix simil	sequence
I-box	Pisum sativum	93	-	0.857	0.883	aACATga
I-box	Pisum sativum	162	+	0.857	0.883	cACATaa
I-box	Solanum tuberosum	163	-	1.000	1.000	tATTATgt
I-box	Pisum sativum	237	-	0.857	0.941	gAAATaa
I-box	Pisum sativum	367	-	1.000	1.000	tATATaa
I-box	Pisum sativum	372	+	1.000	0.941	tATATta
I-box	Pisum sativum	391	-	0.857	0.941	tAAATga
I-box	Pisum sativum	411	-	0.857	0.883	aAAATaa
I-box	Pisum sativum	423	+	0.857	0.883	tAAATta
I-box	Solanum tuberosum	424	-	1.000	0.903	aATAAttt
I-box	Arabidopsis thaliana	426	-	1.000	0.863	aATAAaat
I-box	Arabidopsis thaliana	429	-	1.000	0.863	aATAAaat
I-box	Solanum tuberosum	431	+	1.000	0.951	tATTAttt
I-box	Pisum sativum	433	-	0.857	0.883	aAAATaa
I-box	Pisum sativum	439	-	0.857	0.941	tAAATga

Function: part of a light responsive element

P-box

Site Name	Organism	Position	Strand	Core simil.	Matrix simil	sequence
P-box	Oryza sativa	406	+	1.000	0.857	CCTTttt

Function: gibberellin-responsive element

Prolamin_box

Site Name	Organism	Position	Strand	Core simil.	Matrix simil	sequence
Prolamin-box	oryza sativa	145	+	1.000	0.913	tgaAAAGc
Prolamin-box	Oryza sativa	278	+	1.000	0.913	tgaAAAGc

Function: cis-acting regulatory element associated with GCN4

TATA-box

Site Name	Organism	Position	Strand	Core simil.	Matrix simil	sequence
TATA-box	Daucus carota	79	-	1.000	1.000	TATAaatt
TATA-box	Brassica juncea	80	-	1.000	1.000	TATAaat
TATA-box	Helianthus annuus	81	-	1.000	1.000	TATAaa
TATA-box	Brassica oleracea	82	+	1.000	0.908	tTATAac
TATA-box	Brassica napus	83	-	1.000	0.892	gtTATA
TATA-box	Oryza sativa	117	+	0.818	0.912	TAGAaaa
TATA-box	Oryza sativa	169	+	0.818	0.872	TAAaAAC
TATA-box	Zea mays	248	+	0.909	0.879	TTTAGaaa

Figure 8 (cont'd)

TATA-box	Oryza sativa	250	+	0.818	0.912	TAGAaaa
TATA-box	Oryza sativa	302	+	0.818	0.912	TAAAAaa
TATA-box	Oryza sativa	325	+	0.818	0.912	TAAAAaa
TATA-box	Daucus carota	364	-	1.000	0.863	TATAactt
TATA-box	Brassica juncea	365	-	1.000	0.857	TATAact
TATA-box	Zea mays	366	-	1.000	0.879	TATAtaac
TATA-box	Oryza sativa	367	-	1.000	0.956	TATAtaa
TATA-box	Oryza sativa	368	+	1.000	0.929	TATAtat
TATA-box	Oryza sativa	369	-	1.000	0.929	TATAtat

TATA-box	Solanum tuberosum	370	-	1.000	1.000	TATAta
TATA-box	Glycine max	372	+	1.000	0.891	TATAtt
TATA-box	Oryza sativa	407	-	0.818	0.872	TAAAAag
TATA-box	Zea mays	413	-	0.909	0.879	TTTAAaat
TATA-box	Zea mays	442	+	0.909	0.879	TTTAAaat

Function: core promoter element around -30 of transcription start

TC-rich_repeats

Site Name	Organism	Position	Strand	Core simil.	Matrix simil	sequence
TC-rich_repeats	NT	7	-	1.000	0.952	gTTTTcttca
TC-rich_repeats	NT	68	+	1.000	1.000	aTTTTcttca
TC-rich_repeats	NT	152	-	1.000	0.909	gTTTTcttcg
TC-rich_repeats	NT	191	-	1.000	0.885	tTTTTcttga
TC-rich_repeats	NT	248	-	1.000	0.914	tTTTTctaaa
TC-rich_repeats	NT	285	-	1.000	0.909	gTTTTcttcg
TC-rich_repeats	NT	346	+	1.00D	0.915	aTTTTcttgg

Function:

WUN-motif

Site Name	Organism	Position	Strand	Core simil.	Matrix simil	sequence
WUN-motif oleracea	Brassica	18	-	1.000	0.948	tCATTacat
WUN-motif oleracea	Brassica	139	-	1.000	1.000	tCATTacca
WUN-motif oleracea	Brassica	237	+	0.857	0.948	tTATTtcga
WUN-motif oleracea	Brassica	242	-	1.000	1.000	aAATTtcga
WUN-motif oleracea	Brassica	272	-	1.000	0.948	tCATTacta
WUN-motif oleracea	Brassica	296	-	0.857	0.948	tTATTacgt

Function: wound-responsive element

Figure 8 (cont'd)